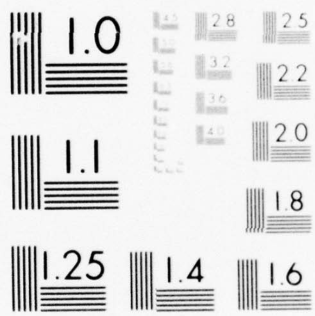
AD-A045 978 ILLINOIS UNIV AT URBANA-CHAMPAIGN DEPT OF CIVIL ENGIN--ETC F/G 8/13
AUTOMATIC SOIL CLASSIFICATION USING QUANTITATIVE TERRAIN FACTOR--ETC(U)
JUL 77 K W WONG DAHC04-74-G-0135
UNCLASSIFIED UILU-ENG-77-2014 ARO-11597.2-GS NL

| OF |
AD
A045978

END
DATE
FILMED
12-77
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ARO-11597.2-GS

UILU-ENG-77-2014

AUTOMATIC SOIL CLASSIFICATION USING QUANTITATIVE TERRAIN FACTORS

Final Report

by

Dr. Kam W. Wong

July 20, 1977

U.S. Army Research Office

Grant No. DAAG29-74-G-0135

Department of Civil Engineering

University of Illinois at Urbana-Champaign

Urbana, Illinois 61801

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 14 UILU-ENG-77-2014	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 6 Automatic Soil Classification Using Quantitative Terrain Factors.		5. TYPE OF REPORT & PERIOD COVERED 9 Final Report, 15 Mar[redacted]74 [redacted] 20 Jul[redacted] 1977
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 Kam W. Wong		8. CONTRACT OR GRANT NUMBER(s) 15 DAHC04-74-G-0135 new DAAG29-74-G-0135
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Civil Engineering, University of Illinois at Urbana-Champaign. Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 16 1T161102BH57 17 01 Atmospheric & Terrestrial
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE 11 20 Jul[redacted]977
		13. NUMBER OF PAGES 30
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 18 ARO 19 11597.2-GS 12 34 p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Soil classification, soil mapping, automatic, terrain factors, multivariate statistics.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A success rate of 60 to 80% was achieved in identifying the soil parent materials of sample areas from among several highly similar parent materials. Only quantitative terrain factors, which were computed from digital elevation and drainage data, were used for discriminating the soil parent materials from each other. The results clearly demonstrated that a reliable method of automatic soil mapping is now technically feasible. over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified 176 010
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Nine soil associations from the state of Illinois and one each from the states of Indiana, Nebraska and Kentucky were used in the study. Eleven commonly used terrain factors were modified for efficient computation using electronic computer, and standard univariate and multivariate analysis techniques were used for testing and classification. Drainage density and surface variance, which is a statistical measure of surface relief, were found to be the most efficient factors for discriminating the soil parent materials.

TABLE OF CONTENTS

Page

List of Illustrations and Tables . . . i

1. Statement of the Problem 1

2. Summary of Most Important Results 4

   2.1 Experimental Data. 4

   2.2 Terrain Factors 6

   2.3 Separability of Soil Associations by Terrain Factors . . 12

   2.4 Accuracy of Automatic Soil Classification. 12

   2.5 Conclusions. 18

3. Publications and Reports 24

4. Scientific Personnel Supported by this Project and Degrees Awarded. 25

5. Grant Numbers. 26

6. Bibliography 27

LIST OF ILLUSTRATIONS AND TABLES

Figure 1. Directions of Traverses in Computing Average Slope, Mean Slope Direction Change and Mean Valley Depth.

Figure 2. Multistage Classification for Illinois Grouping (Quadratic Decision Rule).

Figure 3. Multistage Classification for Combined Grouping (Quadratic Decision Rule).

Table 1. Soil Associations Included in Study.

Table 2. Average Efficiency Indices for Terrain Factors.

Table 3. Significant Differences Between Soil Associations in Illinois Grouping.

Table 4. Significant Differences Between Soil Associations in Combined Grouping.

Table 5. Success Rate in Classification.

1. STATEMENT OF THE PROBLEM

The goal of this project was to develop the basic methodology for automatic soil classification using quantitative terrain factors. The specific objectives were:

1. Develop efficient quantitative terrain factors for soil classification and devise statistical procedure for computing these factors from digital terrain data;
2. Determine the accuracy with which these quantitative terrain factors can differentiate some major soil parent materials found within the glaciated portions in the North Central Region of the United States; and
3. Develop statistical decision logic for soil prediction from quantitative terrain factors.

The identification of soil types and soil characteristics is a necessary prerequisite to the design, planning and construction of many engineering facilities; including route location for highways, site selection for dams, power plants and airports, routing for heavy land vehicles and the location of construction materials. In addition to the laborious procedure of field inspection and soil sampling, air-photo interpretation techniques have been used extensively for soil mapping.

The characteristics of a soil, and by extension its engineering properties, are chiefly a function of the parent material, topography, vegetation, climate and time. Thus, having a thorough understanding of geomorphology and the natural forces of erosion and weathering such as wind, ice and water;

a soil scientist experienced in air photo interpretation techniques can identify soil characteristics from the land forms, erosion patterns, vegetative cover, land use and the tonal distribution in the air photos. The accuracy of the interpretation will depend chiefly on the technical skill of the interpreter and his unique and instinctive ability to detect, correlate and deduce from the many minute hints present in air photos. Since the techniques of air-photo interpretation are based on the recognition of complex photographic patterns and on qualitative analysis, they do not lend themselves easily to automation.

With the recent development in remote sensing technology, many attempts to identify soil types directly from their multispectoral signature have been reported (Anuta et al, 1971; Cihlar and Protz, 1972; Kristof and Zachary, 1974; and Piech and Walker, 1974). The results from these studies generally pointed to the conclusion that an automatic and reliable soil identification technique cannot be based entirely on multispectral analysis. The accuracy of soil identification using either spectral analysis or microwave radiation technique has been degraded greatly by the data noise caused by vegetative cover, atmospheric condition, instability of the sensors and continuous change in sun angle. Even differences in tillage practices on agricultural fields alter greatly the spectral radiance of the surface soil and confound identification by spectral analysis. Another major reason for the failure of spectral analysis technique is that the surface topography, which is a major soil farming factor, is not used as a characteristic factor r soil identification.

The work of Vadnais (1965) and Philips (1970) at the University of Illinois at Urbana-Champaign were the first attempts directed toward the application of quantitative terrain factors to soil classification, although the development of quantitative terrain factors and analytical technique for terrain description have long been an area of active research (Finsterwalder, 1890; Wentworth, 1930; Wood and Snell, 1960; Evans, 1972 and Speight, 1974). Both Vadnais and Philips used manual methods to sample and measure data points from topographic maps. Yet, in spite of the employment of simple statistical sampling procedure and the small number of data points used for each test area, both Vadnais and Philips reported statistically significant correlation between the computed terrain factors and the soil parent materials.

This project was intended to rigorously test the hypothesis put forth by Vadnais and Philips that the type of soil parent materials in an area can be identified from a set of quantitative terrain factors. Digital terrain data including elevations and drainage information were generated from U.S.G.S. topographic maps and air photos for sample areas in which extensive soils information were also available. Algorithms were developed to compute the terrain factors for each test area by electronic computer. Standard statistical methods were used to test the correlation between terrain factors and soils parent materials; and to classify soils using only quantitative terrain factors.

2. SUMMARY OF MOST IMPORTANT RESULTS

The procedure employed and the results obtained in this study will be reported in detail in a Ph.D. dissertation being prepared by M. A. Khoury, who served as research assistant during the entire duration of this project. The dissertation is expected to be completed by September 1977.

The experimental approach and some of the early results have already been reported in a published article by Wong, Thornburn and Khoury (1977). A second paper is being prepared for publication. Therefore, in this final project report, the experimental approach will only be briefly outlined and only the most important results are summarized.

2.1 Experimental Data

One hundred and forty-four (144) sample cells (i.e. sample areas) representing 12 different soil associations were selected from the north central part of the United States. A soil association is composed of several related soil series which have been developed from similar parent materials and have similar soil color. There were 10 to 13 sample areas representing each of the 12 soil associations.

Table 1 presents a summary of the basic characteristics of the 12 soil associations. Nine (9) of the associations were found in the state of Illinois and included thin loess over young glacial (Wisconsinan) till of slightly different texture (I, J and K), glacial outwash (G and X), thick

TABLE 1

SOIL ASSOCIATIONS INCLUDED IN STUDY

Soil Association	Parent Materials	Vegetation	Surface Color	Classification	
				AASHO	Unified
A (Illinois)	Loess > 4-5 ft thick	Prairie	Dark	A-6	CL
G (Illinois)	Medium textured material 2 to 3-1/2 ft thick on gravel	Prairie	Dark	A-2-4	GP
I (Illinois)	Loess < 3 ft thick on loam till	Prairie	Dark	A-6	CL
J (Illinois)	Med. texture material < 4 ft thick on silty clay loam till	Prairie	Dark	A-6	CL
K (Illinois)	Med. texture material < 4 ft thick on silty clay drift	Prairie	Dark	A-7-6	CL
L (Illinois)	Loess > 4-5 ft thick	Forest or mixed prairie and forest	Med. dark to light	A-4 A-6	ML CL
Q (Illinois)	Loess < 4 ft thick on Illinoian drift	Forest	Light	A-6	CL
R (Illinois)	Loess < 7 ft thick over bedrock residuum (sandstone)	Forest	Light	A-4 A-6	CL ML
X (Illinois)	Sand, find sand, loamy sand, fine sandy loam or loamy fine sand	Varied	Light or dark	A-2-4	SM
SH (Nebraska)	Eolian fine sand	None	Light	A-2-4	SP-SM
LS (Kentucky)	Residuum from Cherty limestone	Forest	Light	A-7	MH CH
SS (Indiana)	Residuum from acid sandstone, siltstone and shale	Forest	Light	A-4	ML

loess with different native vegetation (A and L), thin loess over older glacial (Illinoian) till (Q) and thin loess over bedrock residuum (R). The remaining soil associations included soils which were developed from sandstone and shale in Indiana (SS), soils developed from cherty limestone in Kentucky (LS) and eolian fine sand from Nebraska (SH).

Digital terrain data were generated for all the sample areas using U.S. Geological Survey (U.S.G.S.) 1:24,000 and 1:62,500 topographic maps, and aerial photographs were used to help delineate the small drainage ways which were not shown on the map. The 1:24,000-scale maps were used whenever possible. A sample area measured 10 cm by 10 cm on a 1:24,00-scale map and was equivalent to 1.5 miles by 1.5 miles on the ground. When 1:62,500-scale maps were used, the sample cell measured 5 cm by 5 cm on the map, and covered 1.94 miles by 1.94 miles on the ground. The elevation data for each cell was measured in a regular, 21 point-by-21 point grid pattern. Each drainage way or stream was digitized as a series of point for which the x and y rectangular coordinates were recorded. Thus the digital terrain data consisted of spot elevations and positions of drainage ways.

2.2 Terrain Factors

Based on the experience of Vadnais and Philips and on the results of the extensive literature review, eleven (11) terrain factors were selected for this study. Six of these factors related to the surface geometry of the land, and the remaining five related to surface drainage. These factors were defined as follows:

Factors on surface geometry:

1. Sample relief (SR) is defined as the difference between the highest and lowest elevation in the sample area (feet or meters).
2. Sample variance (SV) is defined as six (6) times the standard deviation of the 441 spot elevations within the sample area (feet or meters).
3. Average slope (AS) is the ratio of the total sum of the absolute elevation differentials along several traverses to the total length of these traverses (in percent). Traverses were made in all possible routes along the easterly, southerly, south-easterly and north-easterly directions, see Fig. 1.
4. Mean slope direction changes (MSDC) is the ratio of the total number of slope curvature reversals along the above defined traverses to the total length of these traverses (No. of slope changes/mile of traverse).
5. Roughness index (RI) is the ratio of the surface area as computed from the spot elevations to its orthogonal projection on the horizontal plane.
6. Elevation relief ratio (ERR) is the ratio of the difference between the average and lowest elevations in the cell to the sample relief.

Factors on drainage features:

7. Drainage density (DD) is the ratio of the total length of streams in the cell to the total area of the cell (miles/square mile or kilometers/square kilometer).

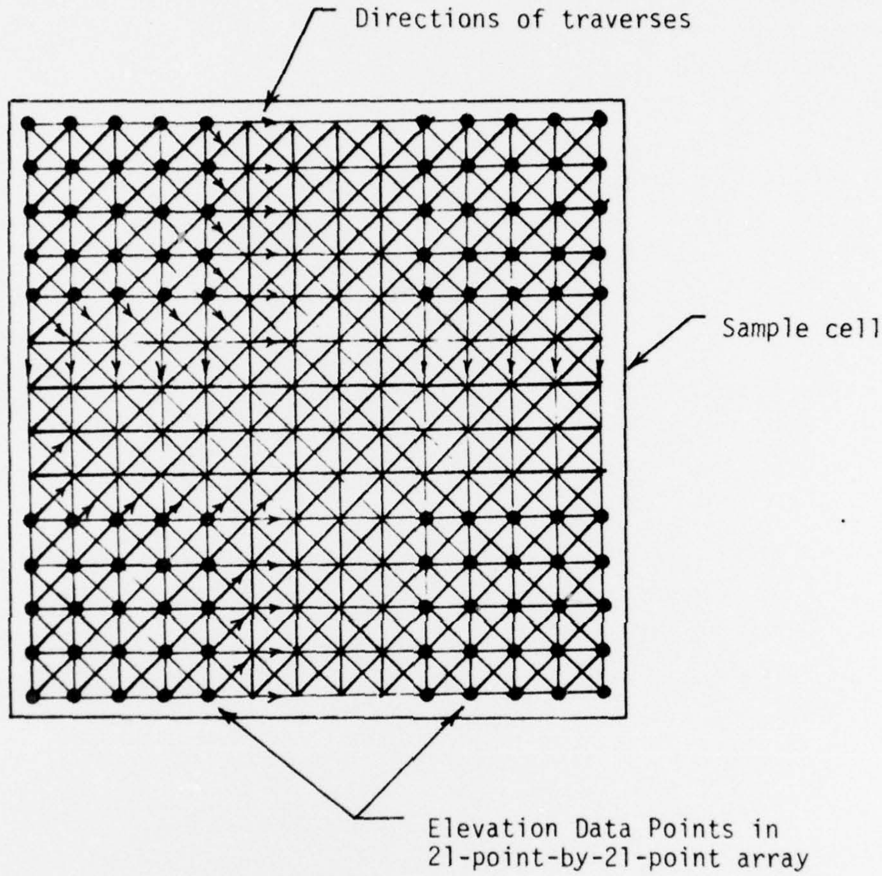


Figure 1. Directions of Traverse in Computing Average Slope, Mean Slope Direction Change and Mean Valley Depth.

8. Ruggedness number (RN) = drainage density x sample relief.

9. Bifurcation angle (BA) is the average value of all junction angles of the streams in the sample cell (degrees).

10. Texture (T) is defined as the ratio of the total number of bifurcations to the total length of streams in the cell (No. of bifurcations/mile or no. of bifurcations/km).

11. Mean valley depth (MVD) is the ratio of the total sum of absolute elevation differentials along the above defined traverses to the total number of slope direction changes along these traverses (feet or meters).

Sample variance and drainage density were found to be the most efficient terrain factors in discriminating the soil parent materials included in this study. The discriminating power of these two factors, as measured by an average α-level of significance of 0.09, is twice as the discriminating power of the two least efficient terrain factors, which were the mean slope direction changes and the elevation relief ratio ($\bar{\alpha} \simeq 0.18$). This finding supports the observation by Strahler (1950) that the two dominant morphometric parameters which control landscape geometry were the relief and drainage density. Table 2 lists the efficiency indices and the rank order of all the terrain factors, except bifurcation angle which was found to have no significant difference among the soil materials investigated.

It should be emphasized that the efficiency ranking given in Table 1 applies only to the soil parent materials and types of topography included in this study. For other types of topography or parent materials, the discriminating power of the terrain factors may be completely different.

TABLE 2

AVERAGE EFFICIENCY INDICES FOR TERRAIN FACTORS

	Factors on Surface Geometry						Factors on Drainage				
	SR	SV	AS	MSC	RI	ERR	DD	RN	BA	T	MVD
Efficiency Indices	0.12	0.09	0.12	0.17	0.15	0.18	0.10	0.13	--	0.11	0.14
Efficiency Ranking	4	1	5	9	8	10	2	6	11	3	7

Both sample variance and sample relief were used to provide a measure of the relief. The main difference between these two factors is that sample relief was computed from the minimum and maximum elevations, whereas sample variance provided a statistical measure of the spread in elevation. It was found that sample variance was more efficient than sample relief in discriminating the soil parent materials found in Illinois.

The average slope was found to be "statistically" equivalent to the sample relief. That is, these two factors were highly correlated, and both factors provided the same basic information in discrimination analysis. Again, this result is in agreement with previous findings by Strahler (1950), Peltier (1954), Wood (1967), King (1968) and Mark (1975) which were obtained under different sets of conditions.

Similarly, the roughness index and mean valley depth were also found to be highly correlated with sample relief.

Largely due to the nature of the topography, the values of the terrain factors were found to be highly variable among sample cells of the same parent materials. The terrain factors on drainage features were twice as variable as those describing surface geometry. The average coefficients of variation for drainage density and surface geometry were found to be 46 percent and 23 percent respectively. On the other hand, the Wisconsinan glacial tills (I, J and K) and outwash materials (G and X) had the most highly variable terrain factors. Their corresponding average coefficients of variation ranged between 31 percent and 62 percent.

2.3 Separability of Soil Associations by Terrain Factors

Pairwise t-test was used to test the ability of a given terrain factor in separating two soil associations. Table 3 summarizes the results of the tests conducted for the nine soil associations found in Illinois. For example, between soil associations A and X, significant differences at the 5 percent level were found in the following terrain factors: elevation relief ratio (denoted by numerial 6), drainage density (7), texture (9) and ruggedness number (10). That is, the values of these terrain factors computed for association A were significantly different from the values of the corresponding factors computed for association X. It can be seen from Table 3 that only associations I and J could not be separated from each other by any of the terrain factors; whereas, between associations I and K and between J and K, only the mean slope change factor was found to be significantly different. All the remaining pairwise combinations of these soil associations could be separated by two or more terrain factors.

Table 4 shows the results of similar pairwise t-tests for another grouping of soil associations which include A, J, X, Q and R from Illinois, SS from Indiana, LS from Kentucky and SH from Nebraska. Any pairwise combination of these terrain factors could be separated by two or more terrain factors.

2.4 Accuracy of Automatic Soil Classification

The primary objective of this study was to determine the accuracy with which quantitative terrain factors could differentiate some major soil parent materials. The test results showed that a success rate of 60 to 80

TABLE 3

SIGNIFICANT DIFFERENCES BETWEEN SOIL ASSOCIATIONS IN ILLINOIS GROUPING*

SA	R	Q	K	J	I	X	G	L
A	1, 2, 3, 4, 5, 7, 10, 11	1, 3, 4, 5, 6, 7, 10, 11	2, 3, 6, 9	6, 9	6, 9	6, 7, 9, 10	2, 6, 7, 9, 10	1, 2, 3, 4, 5, 7, 10, 11
L	1, 2, 4, 5, 10, 11	1, 2, 3, 4, 5, 6, 7	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	
G	1, 2, 3, 4, 5, 7, 9, 10, 11	1, 2, 3, 4, 5, 7, 9, 10, 11	7, 10	2, 7, 10	2, 7, 10	2, 11		
X	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 3, 4, 5, 7, 9, 10, 11	2, 5, 7, 9, 10	7, 9, 10	4, 5, 7, 9, 10			
I	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 3, 4, 5, 7, 9, 10, 11	2, 7, 9, 10, 11					
J	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 3, 4, 5, 7, 9, 10, 11	2					
K	1, 2, 3, 4, 5, 6, 7, 9, 10, 11	1, 2, 3, 4, 5, 7, 9, 10, 11						
Q	1, 2, 3, 4, 5, 6, 7, 10, 11							

* Numbers in each box represent the significant terrain factors between the soil association pair.
Numeric codes for terrain factors: (1) Average Slope, (2) Mean Slope Direction Changes, (3) Roughness Index, (4) Sample Relief, (5) Sample Variance, (6) Elevation Relief Ratio, (7) Drainage Density, (8) Bifurcation Angle, (9) Texture, (10) Ruggedness Number, (11) Mean Valley Depth

TABLE 4

SIGNIFICANT DIFFERENCES BETWEEN SOIL ASSOCIATIONS IN COMBINED GROUPING*

SA	SS	LS	R	X	Q	J	SH
A	1 2 3 4 5 6 7 9 10 11	1 2 3 4 5 6 9 10 11	1 2 3 5 7 10 11	1 2 3 5 7 10 11	1 3 4 5 6 7 10 11	6 9	1 3 4 5 6 7 9 10
SH	1 2 3 4 5 6 7 9 10 11	1 2 3 6 7 9 10	1 2 3 4 6 7 9 10 11	1 2 3 4 5 6 7 9 10 11	1 3 4 5 6 7 9 10 11	1 3 4 5 6 7 9 10 11	
J	1 2 3 4 5 7 9 10 11	1 2 3 4 5 7 9 10 11	1 2 3 4 5 6 7 9 10 11	7 9 10	1 3 4 5 7 9 10 11		
Q	1 2 3 4 5 7 9 10 11	1 2 3 4 5 7 11	1 2 3 4 5 6 7 10 11	1 3 4 5 7 9 10 11			
X	1 2 3 4 5 6 7 9 10 11	1 2 3 4 5 7 9 10 11	1 2 3 4 5 6 7 9 10 11				
R	1 3 4 5 7 9 10 11	2 6 9 11					
LS	1 2 3 4 5 6 7 9 10 11						

* Numbers in each box represent the significant terrain factors between the soil association pair.
Numeric codes for terrain factors: (1) Average Slope, (2) Mean Slope Direction Changes, (3) Roughness Index, (4) Sample Relief, (5) Sample Variance, (6) Elevation Relief Ratio, (7) Drainage Density, (8) Bifurcation Angle, (9) Texture, (10) Ruggedness Number, (11) Mean Valley Depth

percent could be achieved. This is a considerable improvement over the multispectral approach which could yield accuracy of only about 30 to 40 percent (Cihlar and Protz, 1972; Kristof and Zachary, 1974).

Standard methods of multivariate statistics were used to identify the soil parent materials of an area by its terrain factors. For each soil association, a set of training sample cells were used to compute the mean values for the terrain factors as well as a covariance matrix. Then, a test sample cell was classified into the soil association for which its probability of membership is the greatest. The percentage of test cells that were correctly classified then provided a measure of the reliability of classification.

Table 5 summarizes the results of various tests. Again, as in the case of the pairwise t-test, identical tests were made on two groupings of the soil associations. The so-called Illinois grouping consisted of only soil associations which were found in Illinois, i.e. A, G, I, J, K, L, Q, R and X. The combined grouping included associations A, J, Q, R and X from Illinois, SS from Indiana, LS from Kentucky and SH from Nebraska.

Both quadratic and linear decision rules were used. The quadratic decision rule was derived from the assumption that the covariance matrices of the terrain factors for the different soil associations were not equal. The linear decision rule was derived assuming that the covariance matrices of all the soil associations were approximately equal. Although test results showed that there were significant differences (at the 5 percent and 1 percent levels) in the covariance matrices, the linear decision rule was simpler to use

TABLE 5

SUCCESS RATE IN CLASSIFICATION

Grouping	Decision Rules	One Stage Classification: Substitution	One Stage Classification: Leaving-one-out	Multi-stage Classification: Leaving-one-out
Illinois	Quadratic	91%	34%	57%
	Linear	71%	60%	56%
Combined	Quadratic	92%	54%	73%
	Linear	85%	80%	60%

computationally and was therefore included on the analysis. The facts that there was only a small number of sample cells (between 10 and 13) for each soil association, and that some of the terrain factors had zero values for some soil associations (such as drainage density for the sandhills of Nebraska), created some computational problems when quadratic decision rules were used.

Because of the small number of sample cells available for each soil association, three different classification procedures were tested. In the substitution approach, all the available sample cells (10 or 13) for a given soil association were used to compute the mean values of the terrain factors and the associated covariance matrix. Then, each sample was reclassified into one of the soil associations on the grouping. Thus, this approach used the same sample cells as both training and test samples. The best results were obtained with quadratic decision rules, and the success rate amounted to 91 percent and 92 percent for the Illinois and combined grouping respectively.

In the leaving-one-out approach, one of the sample cells was withheld as test samples, while the remaining samples were used to determine the mean terrain factors and the covariance matrix. The withheld sample was then classified into one of the soil associations. This procedure was repeated either 10 or 13 times, depending on the number of samples that were available, for each soil association so that a different sample was withheld as test sample each time. This approach eliminated the build-in bias of the substitution approach, and it made use of the maximum number of samples for computing

the mean values of the terrain factors. Using this procedure, the linear decision rule yielded better results with a success rate of 60 percent for the Illinois grouping and 80 percent for the combined grouping.

The leaving-one-out approach was also used in a multistage classification scheme as shown in Fig. 2 for the Illinois grouping and in Fig. 3 for the combined grouping. For example, in Fig. 2, in the first stage, a test sample was classified as either belonging to group 1 which consisted of L, Q and R, or group 2 which consisted of A, G, I, J, K and X. Only the terrain factors which showed significant differences between these two groups were used in the discrimination analysis. Thus, in this case, only the factors average slope (AS), surface relief (SR), surface variance (SV), mean valley depth (MVD) and ruggedness number (RN) were used. Once correctly classified into a group of soil associations, the test sample was then classified into two smaller subgroups of the group that it was last classified into. This procedure was repeated until the test sample was classified as belonging to one soil association. The quadratic decision rule performed better than the linear rule. It yielded a success rate of 57 percent for the Illinois grouping and 60 percent for the combined grouping.

2.5 Conclusions

This study clearly demonstrated the potential of automatic soils classification using quantitative terrain factors. A success rate of 60 to 80 percent was achieved in identifying the soil parent materials of sample areas from among several highly similar parent materials. It has long been

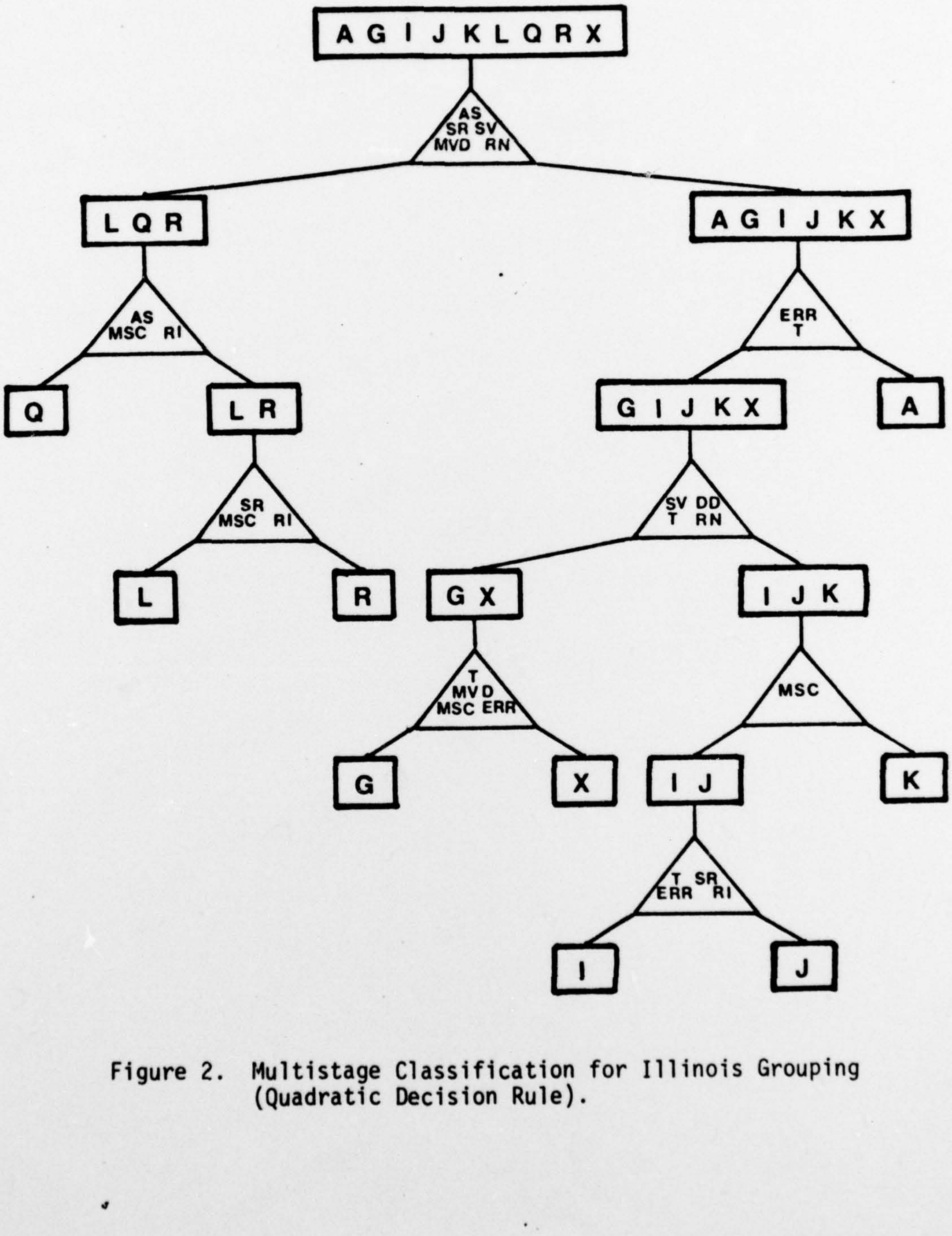


Figure 2. Multistage Classification for Illinois Grouping (Quadratic Decision Rule).

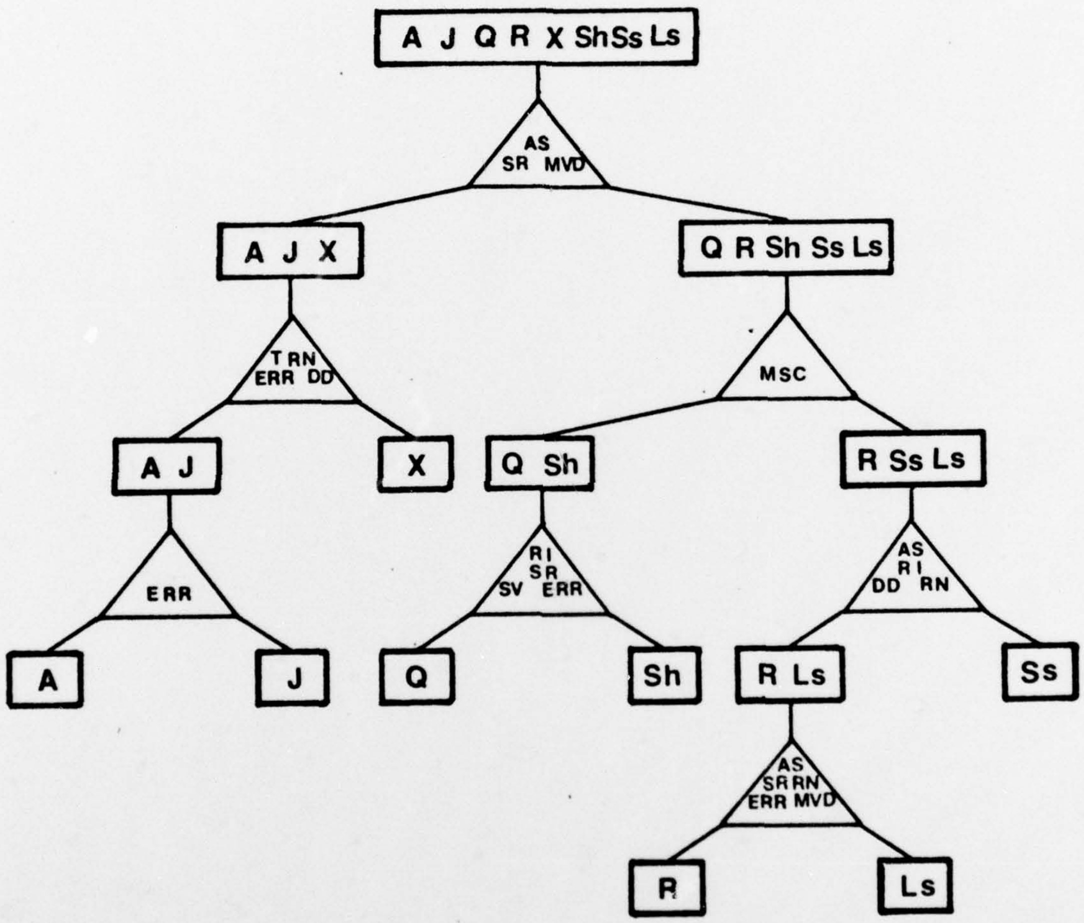


Figure 3. Multistage Classification for Combined Grouping (Quadratic Decision Rule).

recognized that the landforms of an area can provide strong clues to the type of soil parent materials that are present in the area. Geographers and geomorphologists have also long sought to develop simple quantitative parameters to describe the landforms. Vadnais (1965) and Philips (1970) first demonstrated through some simple statistical techniques that quantitative terrain factors could be effective in discriminating soil parent materials. By building on the work of Vadnais and Philips, this study demonstrated that quantitative terrain factors were indeed effective descriptions of soil parent materials. Thus, once the terrain factors have been determined for a group of soil associations, the soil parent materials from unmapped areas can be classified into this group of soil associations with a reasonable degree of accuracy.

Although the tests by substitution method yielded a success rate of 90 percent in soil identification, the approach is highly biased and such reliability can hardly be expected in practice using only terrain factors. It is anticipated that any automatic soil mapping scheme based on quantitative terrain factors will have to make use of other environmental information such as vegetation, climate and land use. These are important soil forming factors and must be considered for any soil mapping scheme to be successful. However, information on vegetation and land use as well as other relevant environmental data can now be collected by remote sensing techniques. Therefore, a reliable method of automatic soil classification using digital terrain and environmental data is technologically feasible. The basic principle of such a method was outlined by Wong, Thornburn and Khoury (1977).

Eleven terrain factors were used in this study. These factors have all been in common use by geographers and geomorphologists. Slight modifications of the definitions were necessary so that these factors could be efficiently computed from digital elevation and drainage data. Some of these terrain factors were very similar to each other, and not all the factors were effective in discriminating any given pair of soil associations. However, it was found that surface variance and drainage density were the most effective factors for the soil associations and topography included in this study. The surface variance is basically a statistical measure which is equivalent to the traditional measure of surface relief. It was found to be a more effective factor than surface relief for the purpose of soil identification.

These conclusions must, of course, be viewed in perspective within the limitation of this study. First of all, only twelve soil associations were included; although these included both highly similar and highly dissimilar associations. For each soil association, only a limited number of sample areas (either 10 or 13) were used in the experiment. Furthermore, each sample area represented a relatively large area, measuring either 1.5 miles by 1.5 miles or 1.9 miles by 1.9 miles depending on whether 1:24,000 or 1:62,500 scale maps were used. There was no attempt made to study the effectiveness of the approach in smaller sample areas. Therefore, the above conclusions are applicable only to soils mapping in a regional scale.

For detailed soil mapping such as that needed for engineering construction, air photo interpretation techniques rely heavily on erosional

characteristics such as gully shapes and on the tonal pattern of the air photo. Since gully shapes are topographic expressions that can be quantitatively measured from large scale topographic maps, and the tonal pattern of the air photo is basically equivalent to data collected by multispectral sensors, it is reasonable to project that the method of automatic soil classification as proposed in this study should be applicable also to detailed soil mapping projects.

This study was confined to areas located within the north central region of the United States. The same parent materials located from different parts of this country or the world may have widely different values of terrain factors. But the primary problem in soil mapping is in discriminating soil materials from within the same region, and selective field testing of the soils can hardly be expected to be completely eliminated regardless of the method used for mapping. Thus, the basic techniques of automatic soil classification should be applicable on a worldwide basis.

Finally, only a few tests were made in applying the cluster analysis techniques for classifying soils. In this approach, all the sample areas were considered as a group and then subdivided into subgroups having similar values of terrain factors. These subgroups were called clusters. Such a technique would be useful for mapping the soils of an area which has little or no existing soil information. The results from these cluster analyses were encouraging, but the number of tests were too few for drawing valid conclusions.

3. PUBLICATIONS AND REPORTS

1. Wong, K. W., T. H. Thornburn and M. A. Khoury, "Automatic Soil Identification from Remote Sensing Data," Photogrammetric Engineering and Remote Sensing, Vol. 43, No. 1, January 1977, pp. 73-80.

2. Khoury, M. A., "Automatic Identification of Soil Parent Material Using Quantitative Terrain Factors," Ph.D. Thesis, Department of Civil Engineering, University of Illinois at Urbana-Champaign, 1977 (in preparation).

4. SCIENTIFIC PERSONNEL SUPPORTED BY THIS PROJECT AND DEGREES AWARDED

1. Dr. K. W. Wong, Associate Professor of Civil Engineering (1971-76), Professor of Civil Engineering (1976-present). Project Co-Investigator from March 1974 to July 1977.

2. Dr. T. H. Thornburn, Professor of Civil Engineering (up to 1975), Emeritus Professor of Civil Engineering (1975-present). Project Co-Investigator from March 1974 to August 1975.

3. Dr. R. Heuer, Associate Professor of Civil Engineering, Project Co-Investigator from August 1975 to July 1977.

4. Mr. M. A. Khoury, Research Assistant on this project from March 1974 to July 1977. Ph.D. degree expected by December 1977.

5. Mr. R. R. Kairam, Research Assistant on this project from May - July 1976.

5. GRANTS OR CONTRACT NUMBERS

1. DAHC04-74-G-0135 (March 1974-January 1976).

2. DAAG29-74-G-0135 (January 1976-July 1977).

6. BIBLIOGRAPHY

1. Anuta, P. E., et al, "Crop, Soil and Geological Mapping from Digitized Multispectral Satellite Photography," Proceedings of the Seventh Symposium on Remote Sensing of Environment, Ann Arbor, Michigan, May 17-21, 1971, pp. 1983-2016.

2. Cihlar, J. and R. Protz, "Color-Film Densities for Soils P. I.," Photogrammetric Engineering, Vol. XXXVIII, No. 11, November 1972, pp. 1091-1098.

3. Evans, I. S., "General Geomorphometry, Derivatives of Altitude and Descriptive Statistics," in Spatial Analysis in Geomorphology, Harper and Rowe, New York, 1972, pp. 17-90.

4. Finsterwalder, S., "Uber Mittleren Boschungewinkel und Das Wahre Areal Einer Topographischen Flache," Stisber, K. Ak. der Wiss., Math.-Phys., Kl., V. 20, 1890, p. 35-82.

5. Khoury, M. A., "Automatic Identification of Soil Parent Materials Using Quantitative Terrain Factors", Ph.D. Thesis, Department of Civil Engineering, University of Illinois at Urbana-Champaign, Illinois, 1977.

6. King. C.A.M., "An Example of Factor Analysis Applied to Geomorphological Data for Six Areas in Northern England," in J. P. Cole and C.A.M. King, Quantitative Geography, Wiley of London, 1968, pp. 319-334.

7. Kristof, S. J. and A. L. Zachary, "Mapping Soil Features from Multispectral Scanner Data," Photogrammetric Engineering, Vol. XL, No. 12, December 1974, pp. 1427-1434.

8. Mark, D. M., "Geomorphometric Parameters: A Review and Evaluation," Geografiska Annaler, No. 57, 1975, A 3-4.

9. Peltier, L. C., "Some Properties of the Average Topographic Slope," Annales Assoc. American Geographers, V. 44 (2), 1954, pp. 229-230.

10. Philips, M. L., "Quantitative Differentiation of Soil Association Areas Using Aerial Photographs," Ph.D. Thesis, Department of Civil Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, 1970.

11. Piech, K. R. and J. E. Walker, "Interpretation of Soils," Photogrammetric Engineering, Vol. XL, No. 1, January 1974, pp. 87-94.

12. Strahler, A. N., "Equilibrium Theory of Erosional Slopes Approached by Frequency Distribution Analysis," American Journal of Science, Vol. 248 (2), 1950, pp. 673, 696; 800-814.

13. Speight, J. G., "A Parametric Approach to Landform Regions," in Progress in Geomorphology, Papers in Honor of David L. Linton, Alden Press, Oxford, England, 1974, pp. 213-230.

14. Vadnais, R. R., Jr., "Quantitative Terrain Factors as Related to Soil Parent Materials and Their Engineering Classification," Civil Engineering Studies, Soil Mechanics Series No. 10, University of Illinois at Urbana-Champaign, Urbana, Illinois, June 1965.

15. Wentworth, C. K., "A Simplified Method of Determining the Average Slope of Land Surfaces," American Journal of Science, Vol. 220, 1930, p. 184.

16. Wood, W. F. and J. B. Snell, "A Quantitative System for Classifying Landforms," Quartermaster Research and Engineering Command, U.S. Army Technical Report EP-124, 1960.

17. Wood, W. F., "Quantitative Considerations in Quantitative Physical Geography," in W. L. Garrison and D. F. Marbe, eds. Quantitative Geography, Part 2, Northwestern University Studies in Geography, 1967, pp. 227-242.

18. Wong, K. W., T. H. Thornburn and M. A. Khoury, "Automatic Soil Identification from Remote Sensing Data," Photogrammetric Engineering and Remote Sensing, Vol. 43, No. 1, January 1977, pp. 73-80.